**ATTACHMENT B**
**PANEL COMMENTS ON JUNE 2009 QPR**
**AND THE EXPERT REVIEW MEETING**

**NCHRP Project 22-24**
*Guidelines for Verification and Validation of Crash Simulations used in Roadside Safety Applications*

Panel comments on the June 2009 QPR are included below in a regular font. The research team's responses are shown in an italic font. Also included are some comments received directly from panel members after the September 9th expert and panel meeting held in Washington, D.C.

**Reviewer #1**

I have read Malcolm's quarterly report and have the following comments:

1. I had a chance to try out the RSVVP program and feel that it will be a useful addition the simulation community. The program is fairly straightforward to use to get results. Having said that, I am a little worried that a lot of effort is going into making this program, rather than focusing on the ultimate methodology that is need for V&V. Although there are advantages to having everyone using the same method, there are also some disadvantages to having a piece of software that is dedicated to this task.
    a. At some point this research project will end and there will no longer be someone to support this piece of software. As computer operating systems change over the years, there may be issues with getting the program to run.
    b. By using software to do much of the V&V, you have a "closed system," whose inner workings are not necessarily changeable or completely understood. This is O.K., as long as the software is doing what you want. However if there is a "bug" or if one wishes to change a piece of the code, this becomes challenging, especially if there is no designated group who maintains the software.
    c. The final report that is issued with this project needs to include the actual methodologies and algorithms, so that these are preserved and can be reviewed and updated as necessary.

    *We certainly agree that there is a certain risk in using software – things certainly change in the software and computer world. That said, however, we also think that a software tool is probably the best way to get people to use the method since it minimizes the effort they have to invest to get the analysis done. One advantage to piggy-backing the program on the MatLab software is that MatLab should take care of most of the long-term software issues and changing the source code will always be upwardly compatible with new versions of MatLab. We certainly intend to include*

*both a User's Manual and a Programmer's Manual in the final report. We have provided drafts of these documents in earlier QPRs. The methods (i.e., the bases of the metrics etc.) will be documented such that a user who wants or needs to do the assessment manual can do so.*

2. In the Benchmark cases, I am somewhat troubled by the use of the Geo Metro model to simulate an impact by a Peugeot, even though they have similar inertial properties (although I realize there may be a limited supply of crash tests to take advantage of.) The issues I see are twofold. If you are successful in your validation, the result suggests that vehicle type (and by extension material properties and etc.) does not matter, as long as the V&V is successful. This may in fact be the case, but this issue may be more appropriately addressed in a separate study. On the other hand, if you are unable to validate the Geo Metro simulations, is this because the event and simulation are truly different, or is your evaluation criteria just too stringent? I guess what I am saying is that it seems you have too many variables. First the test truly is not the same as the simulation (something that should be evaluated more at a later date), and second, we are still working to find the correct way to validate test cases.

   *We certainly appreciate this point. Actually, this is a basic philosophical issue not only with the V&V procedure but our crash testing procedures as well. The same can be said for the V&V exercises we have done using the NCAC C2500 pickup truck model – many of the baseline crash tests did not use that particular vehicle. We have set the acceptance criteria such that we can get passing results for series of crash test (i.e., crash tests NOT simulations) that are considered "equivalent" even when the vehicle was not the same. The ROBUST repeated crash test series did not all use the same make/model/year of vehicles so there is a certain about of "wiggle room" already in the evaluation. The only way to really address this problem is to have some organization fund a repeated series of crash tests with exactly the same vehicle that is available as an FE model. The new Silverado crash tests and simulations help in this respect but this is certainly a good point to keep in mind as we go forward.*

3. The research group is still evaluating acceptance criteria and I would like to make the comment that emphasis should be placed on keeping the evaluation criteria as simple as possible. Results should be easy to understand, not only by the entity submitting the results, but also the FHWA, and others within the roadside safety community. For instance the QPR details the ability to use local or global coordinates and also resultant accelerations. Although this is useful for the team to understand the best way to analyze the data, ultimately one of these methods (ideally the simplest) needs to chosen as the standard for all analyses to be run.

   *We agree. In our past QPRs we have reported on a number of attempts at acceptance criteria but we have begun reducing them to the most relevant and easy to understand results. For example, we no longer make any transformations to the data (e.g., local data is collected in the test and compared to the equivalent local channel in the simulation). Similarly we started with four weighting schemes to look at multi-channel data but now we have eliminated all but the one that we feel gives the best results with the simplest input. We expect retaining the*

*Sprague-Geers M and P components but have eliminated the comprehensive (the C) because it does not provide any additional information. Likewise we expect to retain the ANOVA metrics but have eliminated the T-statistic because it ends up being redundant with the other two ANOVA metrics.*

**Reviewer #2**

In addition to inquiries I sent several weeks ago, I have the following comments about the latest progress report.

1. Acceptance Criteria – I don't really like the "peak values" option. Peak values are very much sampling and filter rate dependent and, I believe, could skew the results too much. The "inertia properties" would be a pain and might be meaningless depending upon type of data curve one is inputting.

   *Neither did we. This has been eliminated.*

2. The report seems to concentrate on acceleration traces of the vehicle too much. There are other data curves that will be used in RSSVP: such as load cell data, accelerometers attached to hardware devices, strain gauges, string pots, etc. This should be clear in both the text (reports, manuals, etc) as well as in the RSSVP menus and results.

   *True. We have been working most on the full crash test comparisons but the other part of RSVVP is certainly important as well. We think this will be made more clear as we continue work on our vehicle and roadside hardware PIRTS since it is the development of the PIRTS where non-accelerometer data is generally used.*

3. Multi-channel seems to be for only the 6 channels from an accelerometer – 3 translation, 3 rotation. I guess I thought it meant I could put in as many or few curves I wanted from various types of singles and then it would give me a general overall view of the correlation. For example, I tried to put in just 2 curves (1) acc of vehicle CG vs time and (2) internal energy vs time, from two separate simulation cases. In RSVVP it prompted me for the next, or 3$^{rd}$ curve, without the option to proceed.

   *No, it is important to remember that this part of RSVVP is specific to comparing crash test signals so this part of the program assumes that the channels represent the various channels used in a crash test. The weighting really makes no sense otherwise. Some improvements have been made, however, to allow users to enter less than six channels in recognition that some labs do not always use all six channels. The latest version of RSVVP allows you to skip some of the channels but in all cases there is assumption that the channels represent accelerometers and rate gyros in a crash test.*

4. All of the test cases do not have to end up having good correlation. It helps to have "bad" cases so that things to look for that are bad is often more helpful than finding things that correlate. But, I understand that people supplying the test cases do not want their examples to come off looking bad in public.

   *We agree and, in fact, some of our test cases have not come out that well. The Geometro comparisons to the ROBUST test, for example, are not as good as we like. Some of the other test cases where we have assessed the models we have been able to document changes that we made to the model that helped the scores. We plan to include all these in the final report.*

5. It sounds like you are making every effort to make available the exact dyna decks and test results being used for the test cases. I support this and think that that is very important. People will want to run the cases themselves as they learn about V & V. That helps reinforce what they'll read about in the reports and give them the training they need to perform V & V on their own with confidence.

   *We agree. Hopefully we can work out an arrangement with the NCAC or FHWA to post the benchmark cases such that the LSDYNA decks are available to people in the future to do both validation and verification checks.*

6. I sometimes get negative values for the magnitude (M) in Sprauge-Geers MPC. That's confusing to me and I could not find any place what a negative value means for magnitude.

   *The negative values in the M metrics just reflect that the simulation magnitude is tending to under-shoot the true curve. The latest version of RSVVP reports the absolute value, however, in order to avoid confusion.*

**Reviewer #3**

Task 8A: Roadside Safety Simulation Validation Program (RSVVP)

This section states, "In a future release of the program, this configuration file may be improved in order to store all of the information necessary to allow a complete reproduction of a previous computer run". I strongly encourage the researchers to make this improvement. The intended use of this program is to support submissions for acceptance of design changes to safety hardware. Therefore, it is important for the approving authority to be able to reproduce the computer run.

*We agree. This should be implemented in the next version of RSVVP (i.e., version 1.7).*

Task 8D: Roadside Safety Model Best Practices Guide

To date, a formal meeting of the panel for NCHRP 22-24 has not been scheduled for this summer in San Antonio. Since NCHRP has not offered to fund travel to San Antonio for the panel members, I do not plan to attend.

I have taken the position that the researcher's earlier proposal to develop a more automated documentation system for LSDYNA models should be tabled until near the end of this project. However, it has become evident that funding from the States for add-on projects will become very tight. An automated documentation system is an interesting tangent, but it is not necessary for successfully accomplishing the objectives of this project. Attachment A shows there is barely enough money remaining to complete this project. Let's use that money to successfully complete the research work that's already underway.

*We have shelved the automated documentation effort.*

Contractual

It is critically important to do the benchmark cases well because this work will give insights into such questions as weighting factors for the PIRTS etc. This has taken more time than originally envisioned and this work is not finished. Therefore, it seems reasonable for the researchers to request a time extension. I recommend that the researchers prepare a revised bar chart schedule and submit a written request for a time extension to NCHRP for consideration by the panel.

*As mentioned in the body of the QPR, the time extension was in fact granted and the new end date is in March 2010.*

**Reviewer #4**

I was impressed by the progress your team has made and feel the research is progressing well. I also thought the inclusion of outside experts was well done. The panel originally envisioned this step as being a presentation of the research with a review by the experts. The way you implemented this step, the experts actually used to the software to analyze test cases, which made their review much more thorough. Some specific comments I have on the presentations and the research to date follow.

1. I though some more about the filtering that users are able to do in RSVVP and believe that the procedures should set a minimum value that is acceptable. For acceleration data, filtering should be done at the frequency of interest, which for NCHRP 350 would be equivalent to a 10ms moving average. If a higher frequency is chosen and the model passes the validation, then it can be adequately used for '350 determination. This means that the model has the necessary bandwidth to accurately reproduce the phenomena that is being simulated. 50ms moving averages are not appropriate, as a model that passes at this level of

filtering may not have sufficient bandwidth to accurately reproduce the phenomena of interest, which is whether the model can pass a test without small 10ms pulses rising above the 20g limit. True, the validation may show that the simulation is reproducing the gross vehicular motion, but it is the 10ms moving average peaks that determine pass or fail in '350. If the simulation is not able to accurately reproduce these, then the pass/fail criteria can not be accurately judged.

*I suppose we have not been explicit in this point but the reviewer is explaining what our intent is. It is our intent that the simulation and test data for a crash test be filtered and compared using the SAE J211 Class 180 filter. This is what Report 350 specifies so the comparison ought to pass at at least this filter level. We will make this clearer in the report and forms. RSVVP allows the user to use a variety of filters and we will probably retain that feature while using a "default" Class 180 filter. We need to include a place in the V&V report forms where the filter class is identified.*

2. I believe the PIRTs are a good idea and are a necessary part of the validation, as they document what the model has been validated for. I do not believe that a rigid set of guidelines should be used as every model (especially hardware models) will be different and will require a different set of validations/calibrations. This will make the review process harder for the "decision makers", as there is no set list of criteria to check against. Still, an innovative piece of hardware that does not fit any of the existing categories should not be unfairly penalized because the designer needs to check "does not apply" to a rigid set of predetermined guidelines. In this case, of course, the designed would need to devise new PIRT criteria that are appropriate to the new hardware design.

   *We agree. We think that the value of the PIRT is that it describes the phenomena such that a decision maker at least knows what is and is not included in the model. It is not necessary for the decision maker to know <u>how</u> the phenomenon is modeled but it at least communicates the modeler's intentions. We too are concerned that a rigid framework for the PIRTs will make things difficult.*

3. After thinking over the comments in the panel meeting, I agree that failed crash tests can be used <u>as part</u> of a validation. In fact, failure is necessary at some level to be able to accurately predict failure. For instance to predict guardrail rupture, you need to have tested the guardrail (or maybe a guardrail coupon) to failure and created an appropriate failure model. So a successfully validated model may have a variety of tests that make up its validation, including some that have failure. At the same time, a successful test cannot be accurately predicted, without a validation that includes a successful test. The original goal of this research was to take a product that had already been certified (i.e. has successful tests), make some small changes, and then run a simulation to verify that the hardware can still successfully perform its function. In other words, the design change was small enough that the change in crash performance is not significant. Taking a test that

has failed (perhaps catastrophically) and then proving through simulation that it no longer fails is a significant change in behavior. For instance using the example discussed during our meeting, a guardrail test that fails due to guardrail rupture cannot be considered successful if design changes are made in the simulation that prevent the rupture. True, the rail may no longer rupture, but does the vehicle now roll? Without a successful test to compare against, we don't know.

*There are a couple of points in this comment. Certainly, the objective of the project had to do with minor incremental changes to existing (presumably accepted) hardware and, as such, there would be no failed crash test. I think where the issue of failure comes up is that often in a development program there are tests where there is failure. So, the developer may have some failed test, a retro-fit and then a new test which did not fail. If all these are available to the modeler, a validation based on <u>both</u> the failed test and the successful test provides much more confidence in the model. Unfortunately, we generally have limited numbers of crash tests to use as baseline times. The point is that we should use as much of the crash test information as is available to get the most confidence. As an example, Ray and Plaxico did some development work on the G2 guardrail for PennDOT. They performed FE simulations to compare against some crash tests where there was guardrail rupture. They then used the model to explore design alternatives that would avoid the rupture. Then they performed a crash test to prove that the FE simulation of the new design was correct.*

**Reviewer #5**

At our last panel meeting, Dr. Ray said that he would accept additional comments until the end of the month. I have the following additional comments…

I suggest that the name of the computer program be changed to the CURVE program (CUrve Resemblance and Validation Evaluation). The name CURVE can be pronounced easily. This will facilitate presentations and discussions. Internet searches for programs dealing with curves will be able to find it. The present name of the program is essentially unpronounceable and it can be confused with the name of the RSAP program.

*As mentioned in the body of the QPR we have put together a survey monkey survey to get some input on various alternative names for RSVVP. We have included CURVE as one of the options.*

In order for this computer program to be credible, it must be free of "bugs". This also implies that the program must be maintained in the future. Is there any provision in the contract for future maintence of the program?

*We certainly we make our best effort to eliminate any bugs in the program prior to the end of the contract. There is, however, no provision for maintenance of the program after the contract is complete.*

**Reviewer #6**

There was a topic introduced at the 22-24 panel meeting that I meant to comment on but missed the window. Someone mentioned that due to material processing, items in guardrails (or vehicles) can be in a pre-stressed state. I remember seeing a talk at the SAE World Congress in 2001 on this topic, particularly related to getting the stresses from a stamping process. Then, last week, at my SAE SEAT committee meeting a code developer mentioned that their software has the capability to model material processes and use the results as an input into crash simulations.

From the response of others in the room, it sounds like this capability has become fairly universal in the FE tools. You may already be aware of this, but in case not, I thought I'd mention it. I'm sure the ls-dyna support can help with this.

*It is certainly possible to model pre-stresses and rolling stresses using LSDYNA. The problem is getting good data. I personally believe, however, that this is a second or third order effect.*

I also remember hearing that you all have a problem similar to mine that different modelers get different results to the same problem. I got an email today that says the nuclear power plant community has the same problem, called user effects. I have yet to hear any solutions, and the problem only becomes obvious when multiple modelers are asked to run the same scenario which is uncommon in certification. However, I worry about user effects in certification scenarios, particularly because the effect is not obvious. I am curious if you have attempted to factor this into your work. If at some point in the future you evaluate this, I'd like to hear about it.

*This is a good point. At this point I don't think there is anything that we can do. The first step is to agree on and use a common method for V&V. After we have some experience in the V&V process and the objective evaluation of models we could do a study where we ask different analysts to model the same baseline crash and see how they compare.*

Also, wanted to let you know that the ASME committee has nearly completed the end to end example, which includes information about documenting a V&V exercise. Len Schwer is on the working group. If you contact him, he may be able to send you a draft.

*Thanks.*

**Reviewer #7**

Attachment C
Since this presentation will probably be used again in the future, I suggest that the words "crash tested" be added to the opening sentence in order to make the objective clear. The following is my suggested wording. "You want to use LSDYNA to develop a new untested variation or design modification of a crash tested roadside device and submit the results to a State DOT/FHWA for acceptance."

*Done.*

Revised Interim Report

The revised interim report is in really good shape. I just have a few comments.

Since the first page shows, "revision 1.3 March 2009", the "DECEMBER 2007" date could be deleted.

*Done.*

Page 9, line 10, "researches" should be "researchers"

The foot note at the bottom of page 36 could read: "Please note that Geers and Gear are two different people." (At this stage of the development of the interim report, any misspellings of their names should have been caught and corrected by the authors.)

*Done.*

Page 83, last sentence. I would suggest the following wording. The SUT model is intended to be a so-called "bullet model", i.e. a vehicle model with a reduced number of elements. This type of model is useful for some types of studies, but it trades off computational accuracy for increased computational speed. I don't want to over-emphasize the point, but I think that something more needs to be said about bullet models from the V&V standpoint.

*We have modified the sentence to read "The SUT model is intended to be a so-called "bullet" model (i.e., a vehicle model with a reduced number of elements) for computational evaluation of roadside safety hardware." We respectfully disagree with the implication of the rest of the sentence. The sentence suggested by the reviewer implies that bullet models are not accurate or correct. In fact, any model is a "bullet" model. There is never enough detail to completely define the mechanical response of a system. What is important is that there be sufficient detail to capture the phenomena of concern. Whether this can be accomplished with 10,000 elements or a million elements is a modeling decision that will be tested by the V&V process. Recall that the definition of validation uses the phrase "for the intended purpose" so the model needs to be detailed enough to accomplish the objectives of the simulation.*

Page 118, eighth line from the bottom. "verification. It should be kept in mind that for V&V purposes, the curves do not have to be exactly identical. Random experimental error makes it unlikely that even curves from two identical tests will be the same. However, for V&V purposes the curves must be *functionally equivalent.*

*Done.*

Contractual

I believe that the 6 month no-cost time extension proposed by the researchers will be marginal because of the remaining work on the test cases and the best practices guidelines. I won't be surprised if an additional time extension will be required.

*We expect to complete the project as presently scoped by the end of March 2010 without any additional time or funds.*

**Reviewer #8**
My only comment/concern is in Attachment D, Table 1: twice hourglass energy is compared to total initial energy.

Do you have any references that explain this? From my experience and LS-DYNA documentation all I've run across is comparison to internal energy at the same specific time. Also, it is very possible to get lots of hourglass deformations with very low hourglass energy reported in glstat or matsum. Visual inspection with mesh turned-on is important. I have no clue how to make that into a measurement, but it is important.